Studies in Avian Biology No. 34:13–33

THE ABCS OF NEST SURVIVAL: THEORY AND APPLICATION FROM A BIOSTATISTICAL PERSPECTIVE

DENNIS M. HEISEY, TERRY L. SHAFFER, AND GARY C. WHITE

Abstract. We consider how nest-survival studies fit into the theory and methods that have been developed for the biostatistical analysis of survival data. In this framework, the appropriate view of nest failure is that of a continuous time process which may be observed only periodically. The timing of study entry and subsequent observations, as well as assumptions about the underlying continuous time process, uniquely determines the appropriate analysis via the data likelihood. We describe how continuous-time hazard-function models form a natural basis for this approach. Nonparametric and parametric approaches are presented, but we focus primarily on the middle ground of weakly structured approaches and how they can be performed with software such as SAS PROC NLMIXED. The hazard function approach leads to complementary log-log (cloglog) link survival models, also known as discrete proportional-hazards models. We show that cloglog models have a close connection to the logistic-exposure and related models, and hence these models share similar desirable properties. We raise some cautions about the application of random effects, or frailty, models to nest-survival studies, and suggest directions that software development might take.

Key Words: censoring, complementary log-log link, frailty models, hazard function, Kaplan-Meier, left-truncation, Mayfield method, proportional-hazards model, random effects, survival.

EL ABC DE SOBREVIVENCIA DE NIDO: TEORÍA Y APLICACIÓN DESDE UNA PERSECTIVA BIOESTADÍSTICA

Resumen. Consideramos como estudios de sobrevivencia de nido se ajustan a la teoría y métodos que han sido desarrollados para el análisis bioestadístico de datos de sobrevivencia. En este marco, la visión adecuada de fracaso de nido es la de un continuo proceso del tiempo, la cual pudiera ser observada solo periódicamente. La sincronización en la captura del estudio y observaciones subsecuentes, así como suposiciones respecto al proceso de tiempo continuo subyacente, únicamente determina el análisis apropiado vía la probabilidad de los datos. Describimos cómo los modelos continuos de peligro del tiempo forman una base natural para este enfoque. Son presentados enfoques no paramétricos y paramétricos, sin embargo nos enfocamos principalmente en el término medio de enfoques débilmente estructurados, y de cómo estos pueden funcionar con programas computacionales tales como el SAS PROC NLMIXED. El enfoque de función peligrosa dirige a modelos de vínculos de sobrevivencia complementarios log-log (cloglog), también conocidos como modelos discretos proporcionales de peligro. Mostramos que modelos cloglog tienen una conexión cercana a modelos de exposición logística y relacionados, y por lo tanto estos modelos comparten propiedades similares deseadas. Brindamos algunas precauciones acerca de la aplicación de modelos de efectos al azar o de falla, a estudios de sobrevivencia de nido, y sugerimos hacia donde pudiera dirigirse el desarrollo de programas computacionales.

A strong interest in nest survival has resulted in numerous papers on potential analysis methods. Recent papers by Dinsmore et al. (2002), Nur et al. (2004), and Shaffer (2004a) have presented methods for modeling nest survival as functions of continuous and categorical covariates and have spawned questions about how the approaches relate to one another. Rotella et al. (2004) and Shaffer (2004a) showed that the Dinsmore et al. (2002) method (which can be implemented in either program MARK or SAS PROC NLMIXED) and Shaffer's (2004a) method are very similar, but how these approaches relate to the Nur et al. (2004) approach is less obvious. In this paper we provide an overview of biostatistical survival analysis. We show how first principle considerations lead to a new nest-survival analysis method based on the complementary log-log link that has practical and theoretical appeal. We focus on techniques designed for grouped or interval-censored data: continuous-time events that are observed in discrete time. We use SAS software (SAS Institute Inc. 2004) for illustration although other environments could be used as well. We discuss and illustrate how current methods used for modeling nest survival relate to methods used in biostatistical applications.

Survival analysis is the branch of biostatistics that deals with the analysis of times at which events (e.g., deaths) occur, and is sometimes referred to as event time analysis. Bradley Efron, inventor of the bootstrap and a leading figure in statistics, described biostatistical survival

analysis as a wonderful statistical success story (Efron 1995). Time is just a positive random variable, apparently qualitatively no different than say weights, which must also be positive. But no large branch of statistics is devoted exclusively to the analysis of weights—what is so special about event times? The answer is how times are observed, or more accurately, how they are only incompletely observed. For example, the classical survival analysis problem is how to estimate the survival distribution from a sample of subjects in which not all subjects have yet reached death; such subjects are said to be right-censored. All we know about right-censored subjects is that their event times are in the future sometime after their last observation. Information on the failure times of these subjects is incomplete. Although perhaps initially counterintuitive, hatching (or fledging) is actually a censoring event because it prevents the subsequent observation of a nest failure. The goal of survival analysis is to extract the maximum amount of information from incomplete observations, which requires a good way of representing incomplete information.

Biostatistical survival analysis has been a relatively specialized domain that has focused mostly on human medical applications. Although some survival-analysis procedures, such as Kaplan-Meier (Kaplan and Meier 1958) and Cox (1972), are fairly widely known beyond biostatistics, the general breadth of survival analysis is not fully appreciated outside of biostatistics. As we discuss, Kaplan-Meier and Cox approaches are seldom well suited to nest-survival analyses and more specialized procedures are generally needed. Our goal here is to show how most nest survival studies can be handled conveniently within the broad framework of modern biostatistical survival analysis theory.

Events in time, such as nest failures, may be incompletely observed in many ways. Two general mechanisms that occur in most nesting studies are left-truncation (resulting from delayed entry) and censoring (exact failure age unknown). Given the various ways in which observations can be incomplete, how can one be assured that the maximum amount of information is being recovered from each observation? This is where the data-likelihood function is important. A correctly specified data likelihood describes the precise manner in which observations are only partially observed. Loosely speaking, the likelihood principle and the related principle of sufficiency imply that the data-likelihood function captures all of the information contained in a data set (Lindgren 1976). No analysis can be better than one based on a correctly specified likelihood.

The likelihood principle says that the data likelihood is the only thing that matters. In some cases, identical likelihoods arise from apparently very different types of data. For example, likelihoods that arise from event-time data are quite frequently identical to likelihoods that result from discrete-count data. By recognizing such equivalences, it is possible to use software to perform event-time analyses even if the software was originally designed for other applications such as Poisson or logistic regression of discrete-count data (Holford 1980, Efron 1988).

Once the data likelihood is constructed, the rest of the analysis follows more or less automatically. Two factors solely determine the data likelihood: data-collection design, and biological structure. Data-collection design refers to how the data are observed and collected, and determines the macro-structure of the likelihood. Biological structure reflects the assumptions or models the researcher is willing to make or wants to explore with respect to the nest-failure process. Biological assumptions and models are usually formulated in terms of the instantaneous-hazard function, and the hazard function in turn determines the micro-structure of the likelihood. Together, the data collection design and biological structure fully specify the data likelihood which forms the foundation of analysis. The need to correctly construct the appropriate data likelihood does not depend on whether one is taking a Bayesian or classical (maximum likelihood) approach to estimation and inference; both approaches are based on the same data likelihood. Here we focus on the maximum likelihood (ML) method which underlies both the classical frequentist approach as well as the recently popularized information-theoretic approach of Burnham and Anderson (2002). We focus on ML methods primarily because of tradition and readily accessible software.

Once the data are collected, the macro-structure of the likelihood is essentially set. The researcher has little or no discretion with respect to structuring this portion of the likelihood once the data are in hand. From the data-collection design it is usually clear what macro-structure is needed. The only reason to use an analysis that is not based on the exact macro-structure is because it is exceedingly inconvenient. In such cases, researchers can try analyses with likelihood macro-structures corresponding to data-collection designs that they hope are close enough to give good approximations. Mayfield's (1961, 1975) method, including Mayfield logistic regression (Hazler 2004), is an example of an analysis that is based on

an approximate macro-structure as a result of the unrealistic assumption that failure dates are known to the day (i.e., Mayfield's midpoint assumption). Johnson (1979) and Bart and Robson (1982) derived an exact analysis for the problem considered by Mayfield, but these methods have received relatively little use because software was not readily available at the time. Because it is difficult to say when an approximate likelihood is close enough, one should always strive for a likelihood as accurate as possible. The consequences of such assumption violations can range from negligible errors to completely invalid results, affecting both estimation and testing.

The researcher has much more freedom with respect to the biological structure, and this is the aspect of nest-survival analysis that requires some creativity and judgment. In biostatistical survival analyses, so-called nonparametric procedures such as the Kaplan-Meier estimator (KME) and the Cox partial likelihood approach enjoy great popularity because of the perception that they can be applied almost unconsciously on the part of the researcher. However, things are often not so simple with nest-survival data. In fact, many nest-survival data sets cannot support fully nonparametric approaches because of left-truncation and interval-censoring, which will be described later. Indeed, nonparametric is a misnomer; nonparametric survival approaches actually require the estimation of many more parameters than typical parametric analyses (Miller 1983), which is why they are not a panacea in nest-survival studies.

Due to the low data-to-parameter ratio in fully nonparametric procedures, the resulting survival estimates typically have large variances. The primary appeal of fully nonparametric procedures is that under some circumstances the estimates can be counted on to be relatively unbiased and moderately efficient (although left-truncation and interval-censoring, common features of nest survival studies, may result in exceptions; Pan and Chappell 1999, 2002). The situation is reversed for so-called parametric approaches. The survival estimates from parametric survival models typically have small variances because few parameters must be estimated. However, this can be at the price of large biases. In statistics in general, it has long been recognized that the best estimators are those that achieve a balance between variance and bias, which is measured by the mean squared error. Thus, in many survival-analysis situations, including nest survival, the best approach is the middle ground between fully nonparametric approaches and traditional parametric models; this middle ground is often referred to as weakly structured models, which we will explore in the nest-survival context.

Our intention is to present practical ideas that will be useful in the analysis of real data. To facilitate this, we use an example data set throughout the paper to illustrate how particular ideas translate specifically into analyses. All programs used for the analyses are given in the Appendices.

PROBABILITY BASICS

Symbolic Representation of a Nest Record

We will use T to represent the actual age at which a nest fails. In most cases, this quantity will not be observed exactly or at all, but we can always put bounds on it. A nest record needs to describe two things: (1) the age observation starts (discovery), and (2) what bounds we can put on the failure age T. For example, suppose we discover a nest at age r, and follow it until age t. Suppose age t is the last we observed the nest, at which point it was still active. Symbolically, we will describe such a nest observation as $T > t \mid T > r$, which means starting at age r (conditional on being active at r), the nest was observed until age t, and had not yet failed. Another nest, discovered at age r, still active at age x, but failed by age t would be described as $x < T < t \mid T > r$.

Nest Record Probabilities

The data likelihood gives the probability of the observed data. It is constructed by first computing the survival probability (or survival probability density in some cases) corresponding to each nest record, and then multiplying all of these nest-likelihood contributions together. The age of nest failure T is a random variable that is characterized by its probability distribution. For the record described by $T > t \mid T > r$, $\Pr(T > t \mid T > r)$ is its probability. This is the probability of the nest surviving beyond age t conditional on it being active at age r. It is often more convenient to write this using the shorthand $S(t \mid r) = \Pr(T > t \mid T > r)$. A very important special case occurs when the record starts at the origin (nest initiation) $S(t \mid 0) = \Pr(T > t \mid T > 0)$; this is referred to as the survival function, and is often represented as just $S(t)$. The general goal of survival analysis is often to estimate and characterize $S(t)$. Even if one is only interested in an interval survival such as a monthly rate, $S(t)$ is the means to that end; for example, if age is in days, $S(30)$ is the monthly survival rate.

A very fundamental property of conditional survival probabilities is that they multiply. So for

ages $a < b <$ c, then $S(c \mid a) = S(b \mid a)S(c \mid b)$. In particular $S(t) = S(1 \mid 0)S(2 \mid 1)\ldots S(t \mid t-1)$ (of course assuming age t is an integer). The importance of this multiplicative law of conditional survival in survival analysis cannot be overemphasized.

Suppose we discovered a nest at initiation (age 0), and visited it periodically. We observe that it failed between ages x and t. This observation is described as:

$$x < T < t \mid T > 0,$$

and it should seem reasonable that

$$\Pr(x < T < t \mid T > 0) = S(x) - S(t).$$

From the multiplicative law

$$S(t) = S(x)S(t \mid x),$$

so this can also be written as

$$\Pr(x < T < t \mid T > 0) = S(x)(1 - S(t \mid x)).$$

The term $1 - S(t \mid x)$ is especially important in survival analysis, and is referred to as the conditional interval mortality. It is the probability of failing in the age interval x to t, given one starts the interval alive at age x. We can represent this as

$$\Pr(x < T < t \mid T > x) = 1 - S(t \mid x) = M(t \mid x).$$

LIKELIHOODS

Data-collection Designs—Likelihood Macro-structure

Nest-study data-collection designs, which determine the likelihood macro-structure, can be broadly categorized into three general cases, given below. In a certain sense, the macro-structure is not scientifically interesting, although it must be accommodated to get the right answer. It reflects how the data were collected and is not directly influenced by biology. By interval monitoring, we mean that some interval of time elapses between visits to the nest; the inter-visit intervals need not all be of the same duration. If a nest fails, the failure time is known only to have been sometime during that interval. Without going into the details, under continuous monitoring the contribution of a failed nest to the likelihood is technically a probability density rather than a probability per se.

Case I: Known age, continuous monitoring:
- Discovered at age r:
  - Last observed active at age t:
    $\Pr(T > t \mid T > r) = S(t \mid r)$
  - Observed failure at exactly age t:
    $\Pr(t < T < t + dt \mid T > r) \approx S(t \mid r)h(t)dt$;
    $h(t)$ is a hazard function.

Case II: Known age, interval monitoring:
- Discovered at age r:
  - Last observed active at age t:
    $\Pr(T > t \mid T > r) = S(t \mid r)$
  - Observed failure between ages x and t:
    $\Pr(x < T < t \mid T > r) = S(x \mid r)(1 - S(t \mid x))$.

Case III: Unknown age, continuous or interval monitoring:
- Age at discovery known only to be between r_y (youngest possible) and r_o (oldest possible):
  - Last observed active time d after discovery:
    $\Sigma p(r)S(d + r \mid r)$;
    $r_y \le r \le r_o$
    $p(r)$ is the probability of discovery at age r
  - Observed failure between z and d days after discovery ($z < d$)
    $\Sigma p(r)S(z + r \mid r)(1 - S(d + r \mid z + r))$
    $r_y \le r \le r_o$

Case I allows for left-truncation (delayed discovery) and right-censoring (some failures never observed) and is very important in human biomedical applications, but is seldom appropriate in nesting studies. Case II allows for left-truncation, interval-censoring (failure time known only to an interval), and right-censoring. Case III allows for left-truncation and general double-censoring (Heisey and Nordheim 1995). While Case III is the most general, it is not yet straightforward in application due to software issues. We focus most of our attention on Case II—known-age, interval monitoring.

The Geometric Interpretation of Likelihood Contributions

The basics of the macro-structure likelihood contributions become clear by considering the Lexus diagram (Fig. 1). The Lexus diagram has a long history in survival analysis (Anderson et al. 1992), and is extremely useful for visualizing the likelihood contributions in complex situations involving delayed discovery and interval-censoring, especially in the most general case when survival can vary both by age and calendar time, which we briefly consider later. The Lexus diagram displays the known history of a nest in the calendar time/nest age plane. One can imagine a probability density spread over this two-dimensional surface. To determine the likelihood contribution, one has to first determine the region on the time/age plane that is being described by the nest record. One then collects the appropriate probability over this region.

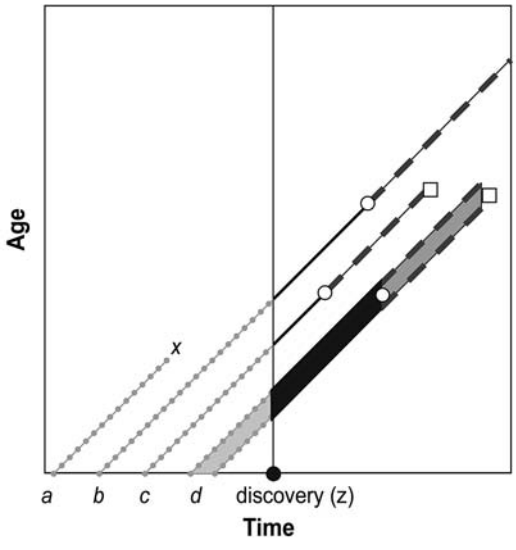

FIGURE 1. Lexus diagram showing some possible observational outcomes for four nests in a typical survival study. The nests are indicated as *a*, *b*, *c*, and *d*. We will also let *a*, *b*, *c*, and *d* indicate the dates of nest initiation. A hollow circle indicates the last visit during which the nest was known to be active, and the hollow square indicates the first visit at which the nest was known to have failed. We assume nests were searched for on only one day, say *z*. Nest *a* is an example of a hypothetical nest that failed before discovery on day *z*, and hence was unobservable (left-truncated). Nests *b* and *c* are examples of nests that were discovered on day *z* and determined to be exactly $z - a$ and $z - b$ days old. Nest *b* went on to hatch, so its hypothetical failure time can be thought as being sometime during the infinite interval after hatching. Nest *c* was observed to fail sometime during the indicated interval. The likelihood contributions mirror this structure. Nest *d* could not be aged exactly, so its date of initiation can only be bounded. Such unknown ages result in a two-dimensional region over which probability density must be collected, which is why Case III likelihood contributions are sums.

The histories of four nests are shown (Fig. 1). For simplicity of illustration, nests were searched for on only one day, labeled discovery on the *x*-axis. The day of discovery is the so-called truncation limit; nests that do not survive until that day are truncated from the potential sample and their existence is never known. Nest *a* is an example of a truncated nest. If we had discovered the remnants of nest *a*, this would constitute a left-censored observation; failure occurred to the left of the first observation. We do not deal with such problematic observations in this paper. Nests *b*, *c*, and *d* are examples of discovered nests. The ages of both nests *b* and *c* were determined exactly at the time of discovery, so their records are known to lie on a line in the time/age plane. The hollow circle indicates the last visit at which the nest was active, and the hollow square indicates the first visit when the nest was known to have failed. The solid line to the right of discovery indicates when the nest is known to have been active, and the broken line is the region in which the nest could have failed. Nest *c* was observed to fail in an interval (say between x and t), after first surviving for an interval from r to x. This history is described as $(x < T < t \mid T > r)$, with corresponding probability:

$$\Pr(x < T < t \mid T > r) = S(x \mid r)(1 - S(t \mid x)).$$

Nest *b* was never observed to fail (right censored), but the geometry of its observation can be viewed in exactly the same manner as nest *c*. We assume nest *b* would hypothetically fail sometime between the last observation and infinity, so its record is $(t < T < \infty \mid T > r)$. The corresponding probability statement is $\Pr(t < T < \infty \mid T > r) = S(t \mid r)(1 - S(\infty \mid t))$. Of course the probability of surviving forever is 0, $S(\infty \mid t) = 0$, so the likelihood contribution for a right-censored observation reduces to $\Pr(T > t \mid T > r) = S(t \mid r)$, as given before. This shows that right-censoring is just a special case of interval-censoring where the upper bound is infinity.

Nest *d* illustrates the case where a nest's age at discovery could only be bounded. The black polygon indicates time/age points when the nest could have been active, and the grey polygon indicates time/age points when the nest could have failed. The Case III likelihood contributions reflect the sums over these two-dimensional regions.

In the Lexus diagram nest age and calendar time are continuous variables. This is realistic; a nest can fail at any time day or night. In almost all cases it is appropriate to think of the event of nest failure as a continuous-time event, even if it is not observed or recorded in continuous time. This continuous-time event framework is the framework on which most of modern biostatistical survival analysis theory rests. Its power lies in its ability to accurately represent how data are incompletely observed under a diversity of circumstances as suggested by the Lexus diagram. Failure to accurately represent the continuous time region in which the observation may have occurred is likely to result in biases. An obvious example of this is the well-known issue of apparent survival versus the Mayfield estimator; Heisey and Nordheim (1990) give a more complex example.

Example

We now introduce an example that we will use throughout this paper for illustration. It is

a sample (N = 216) of Blue-winged Teal (*Anas discors*) nests taken in 1976 reported by Klett and Johnson (1982). Nests in the sample were obtained by searching right-of-way habitat along Interstate 94 in south-central North Dakota. The macro-structure of the data set is classic general Case II—aged nests discovered sometime after initiation with periodic re-visitation (Fig. 2). Few of the nests were discovered on or near the time of initiation, so as suggested by Fig. 2 the data contain very little survival information with respect to the youngest ages. On Fig. 2, a solid black line segment indicates an age span during which it is known that the nest survived. A black segment going from age r to age t contributes the term $\Pr(T > t \mid T > r) = S(t \mid r)$ to the likelihood. A dashed-line segment indicates an age span during which it is known that the nest failed. Such a segment going from age x to t contributes: $\Pr(x < T < t \mid T > x) = 1 - S(t \mid x)$ to the likelihood. These are the correct likelihood contributions for the observational design of the study, and in addition to demonstrating appropriate approaches, one of our goals will be to examine the consequences of using less appropriate analyses.

The data file contains five variables. One variable is the nest identifier nestid. The variables firstday and lastday are the first and last days of a visitation interval; the days on which visits occurred. The variable success indicates whether the subject survived the interval (1) or not (0). The variable distance gives the distance to the road shoulder. A nest often had multiple records, one for each inter-visitation interval. However, no loss of information occurs by combining all consecutive successful intervals for a nest and treating them as a single interval. This follows since: $S(b \mid a)S(c \mid b) = S(c \mid a)$.

Continuous-time Events, Hazard Functions, and the Daily Survival Rate

The hazard function $h(t)$ is the key to representing survival probabilities in continuous time; it is the basic structure on which all else rests in survival analysis. It links the probability surface over the Lexus diagram to interesting biological models. The best way to think of $h(t)$ is as the conditional interval mortality scaled per unit time,

$$h(t) \approx \frac{M(t + dt \mid t)}{dt}$$

i.e., the instantaneous failure rate. It is formally defined as the limit of this relationship as dt goes to 0. Hazard functions are particularly suitable for regression modeling. The hazard function uniquely determines the survival function through the rather opaque relationship:

$$S(t \mid r) = \exp\left[-\int_{r}^{t} h(u)du \right] \quad (1)$$

The specific form of this relationship should be viewed more-or-less as just math; relatively little intuition can be gained from studying it although it is a key mathematical relationship to know. The term

$$\int_{r}^{t} h(u)du$$

is very important in modern survival analysis, and is referred to as the cumulative interval hazard; we will represent it with the more convenient notation

$$\Lambda(t \mid r) = \int_{r}^{t} h(u)du$$

Just as conditional survival probabilities multiply, cumulative interval hazards add: $\Lambda(c \mid a) = \Lambda(b \mid a) + \Lambda(c \mid b)$. This additivity is quite convenient.

Usually nests will not be visited more than once daily and we assume that this is the case in this paper. This is convenient because we can assume age t is always an integer and use the daily cumulative hazard $\Lambda_t = \Lambda(t \mid t - 1)$ as the basic building block and avoid showing integrals almost entirely (i.e., the integral in (1) is replaced by a sum). This now provides a firm theoretical underpinning for the traditional approach of using daily survival rate (DSR) in nest survival analyses. That is, if DSR_t is the daily survival rate for day t, $DSR_t = S(t \mid t - 1) = \exp(-\Lambda_t)$. Thus, the cumulative daily hazard can be viewed as just a one-to-one transformation of the DSR, $\Lambda_t = -ln(DSR_t)$. By recognizing this relationship between the DSR and the cumulative daily hazard, DSR models can be constructed which have clear hazards-based interpretations.

In ordinary regression analysis, we are accustomed to parameters (slopes) having any possible value, negative or positive. But because hazard functions $h(t)$ must be non-negative, cumulative interval hazards such as Λ_t must be non-negative as well. We can get around this range restriction by using the log cumulative daily hazard $\gamma_t = ln(\Lambda_t)$ for modeling. The relationship of the log cumulative daily hazard to the DSR is then:

$$DSR_t = S(t \mid t - 1) = \exp(-\exp(\gamma_t))$$

This can be rewritten as:

$$\gamma_t = ln(-ln(1 - DMR_t))$$

where DMR is the daily mortality rate 1 – DSR.

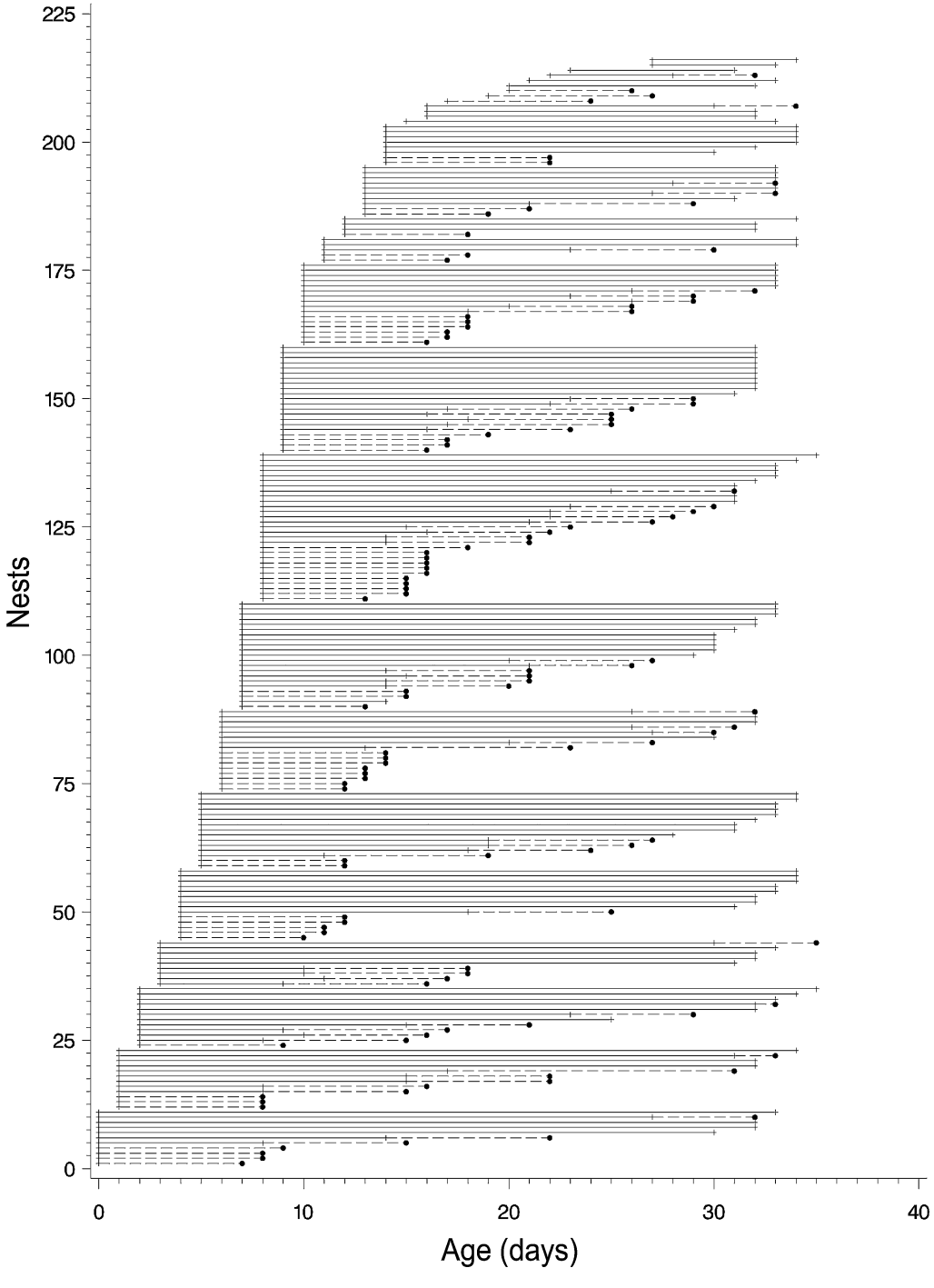

FIGURE 2. Raw data for 216 Blue-winged Teal (*Anas discors*) nests. Solid lines indicate times at which the nest was under observation and known to have survived. Dashed lines ending with a solid dot indicates intervals during which nests are known to have failed.

This important relationship is often referred to as the complementary log-log link model because it links the daily cumulative hazard to the mortality (or survival) function; it is also referred to as the discrete proportional-hazards model. We have been unable to discover with certainty why this model is traditionally given in its complementary form, i.e., in terms of DMR rather than DSR, but without going into the details we believe it is because $ln(-ln\ (1 - P))$ is quite similar to the logit model logit(P), while $ln(-ln\ (P))$ is not. On this scale, we can build familiar-appearing regression models, where the parameters have very clear hazards-based interpretations.

To summarize, for Case II likelihood contributions such as our example, the basic building block is the conditional interval survival, say $S(t \mid r)$. We will assume visits are at the beginning of a day, so visits on days i and j corresponds to the age span $i - 1$ to $j - 1$. Thus, $S(t \mid r) = DSR_{r+1}DSR_{r+2}\ldots DSR_t$. This in turn can be expressed as:

$$S(t \mid r) = \exp(-(\Lambda_{r+1} + \Lambda_{r+2} + \ldots + \Lambda_t)), \quad (2)$$

and $\Lambda_s = \exp(\gamma_s)$.

Equation (2) can be expressed in pseudo-code as:

```
total_cumulative_hazard ← 0
for day = firstday to lastday - 1 do{
  daily_cumulative_hazard ←
    exp(gamma[day])
  total_cumulative_hazard ←
    total_cumulative_hazard +
    daily_cumulative_hazard
}
interval_survival ←
  exp(-total_cumulative_hazard);
```

Any Case II analysis will have this general structure at its core because this general structure accommodates the likelihood macrostructure. Most of the remainder of this paper focuses on various models for the vector gamma, which gives the micro-structure. The importance of (2) in general Case II applications is difficult to over emphasize. (Aside: time indexing for such analyses can be rather confusing. In the above pseudo-code, because visits are assumed to occur at the beginning of the day, the last full day survived is the day before the last visit, hence lastday-1.)

So the total data likelihood is a product of terms of the form $S(t \mid r)$ and $1 - S(t \mid r)$. In this respect, even though the random variable being modeled is actually the continuous variable age at failure, the likelihood appears exactly the same as one that would arise from binary or binomial data. This is very convenient because it allows us to use software intended for the analysis of discrete binary or binomial data. For our examples, we used SAS PROC NLMIXED specifying a binary model.

SURVIVAL ESTIMATION

The Simplest Example—General Case II, Constant Hazard

We start with the simplest (and most restrictive) possible model, which is under the assumption that the hazard does not vary with age, so $h(t) = \lambda$. When applied to general Case II data, this estimator corresponds to the generalization of the Mayfield model developed by Johnson (1979) and Bart and Robson (1982). Under the special circumstance of Case II data resulting from once-daily monitoring, Mayfield estimates are obtained. Under this model, all values of the vector gamma are the same, regardless of age (Program A-1; Appendix 1). The result of applying this model to the example data is shown on Fig. 3. With respect to the hazard function $h(t)$, this is the most restricted and smoothest possible model. With this as background, we next look at the least restricted and roughest possible models with respect to $h(t)$, so-called nonparametric models.

Case I and Special Case II—Nonparametric Survival Estimation

Nonparametric is a somewhat murky term in statistics with multiple meanings. In survival analysis, a nonparametric survival estimator is usually defined as one that converges exactly to the true survival function $S(t)$ as the sample size grows to infinity for any $S(t)$ (Kaplan and Meier 1958). The counterexample is a parametric survival estimator which will converge to the true $S(t)$ only if the true $S(t)$ happens to belong to the specified parametric family. For a nonparametric estimator to converge to $S(t)$ for every possible $S(t)$, such an estimator must be extremely flexible.

From a theoretical standpoint, a big difference exists between truly continuous monitoring (Case I) and almost continuous periodic monitoring (once daily monitoring—Special Case II). Theoretical justification of continuous-monitoring estimators typically involves rather sophisticated theoretical devices—this has to do with the fact that the probability of a continuous random variable ever assuming a specific value is 0. Kaplan and Meier (1958) achieved biostatistical fame primarily because of their

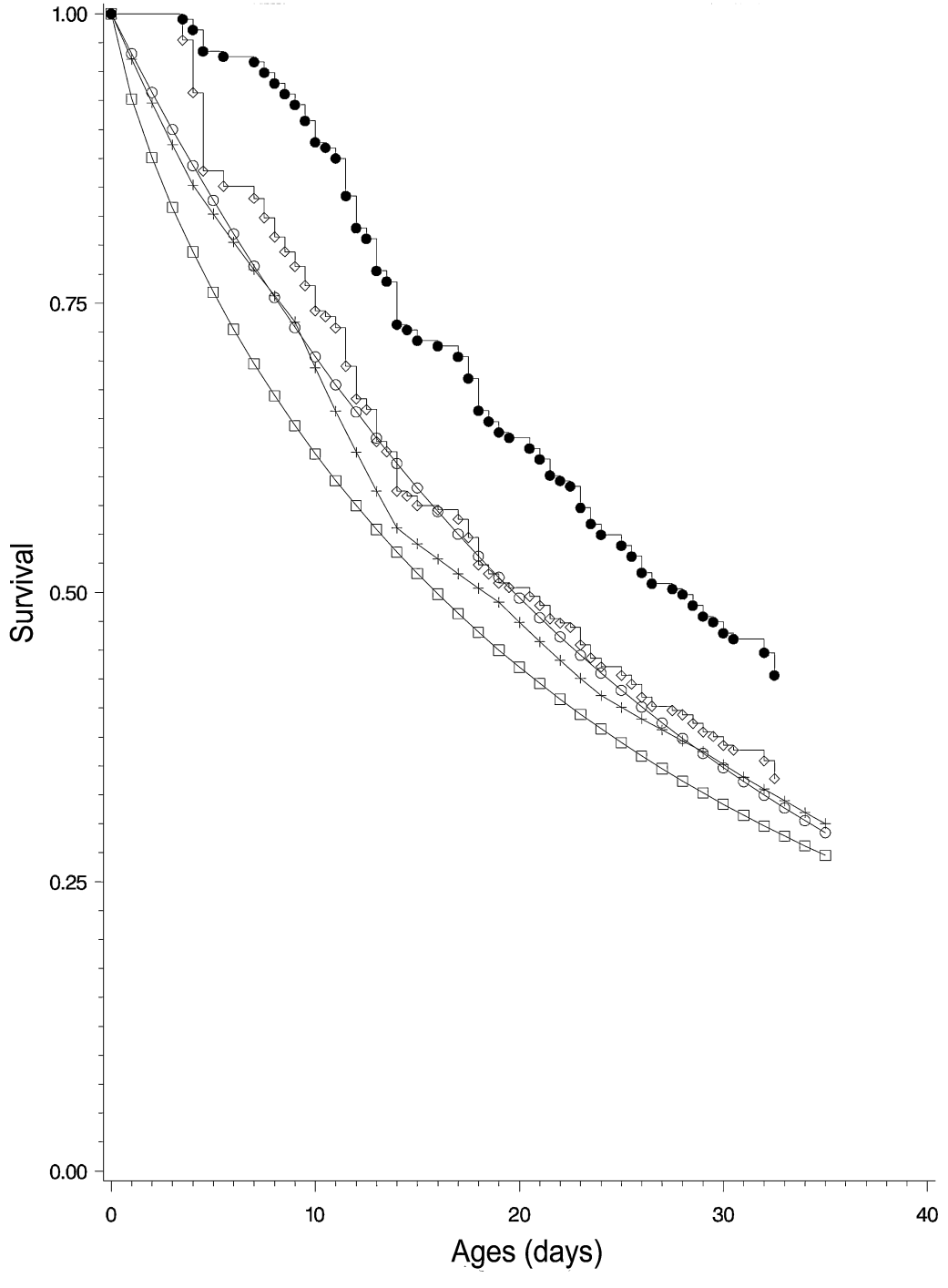

FIGURE 3. Estimated survival curves. The upper most curve (solid dots) is the usual Kaplan-Meier estimator (KME), which ignores the left-truncated (delayed entry) aspect of the data. The generalized Kaplan-Meier estimator (GKME) which accommodates left-truncation but not interval-censoring is the step function with hollow diamonds. The hollow circles correspond to the constant hazard model, the hollow squares to the Weibull model, and the crosses to the weakly structured model with a step-hazard model (steps every 5 d).

clever argument showing that the KME is the nonparametric maximum likelihood estimator (NPMLE) of $S(t)$ specifically under continuous monitoring. In application, this distinction is often not so important—for example, the KME for continuous monitoring and the life table (actuarial) estimator for frequent periodic monitoring are identical, so there seems little harm in referring to both as KMEs as is frequently done. In the following we focus on once-daily monitoring, and occasionally blur the distinction between continuous and once-daily monitoring a little to avoid tedious qualifications.

As noted, for a nonparametric estimator to converge to $S(t)$ for every possible $S(t)$, such an estimator must be extremely flexible. The manner in which nonparametric estimators typically achieve this is by allowing the empirical hazard to change whenever a failure is observed. Two popular approaches are the impulse-hazard model and the step-hazard model.

To justify the impulse-hazard model, it can be argued that it is reasonable to assume that on a day when no failures occur, the cumulative daily hazard Λ_t is 0. But on a day a failure occurs, Λ_t spikes up but then falls back down the next day if no failures occur. Under the step-hazard model, it can be argued that it is reasonable to assume the daily cumulative hazard Λ_t remains constant (and not necessarily 0) until after the next failure occurs, but that it might step up or step down at that point. Both of these models are extremely flexible, perhaps in some sense too flexible.

Either of these hazard models can be implemented relatively easily within our general framework outlined earlier. Let $t_{(1)}, t_{(2)}, \ldots$ indicate the days on which failures were observed. For the impulse-hazard model, the easiest approach is simply to discard any days on which no failures occurred and then allow γ_t to be different for each day $t_{(i)}$ on which failures were observed. To implement the step-hazard model, the γ_t of the gamma vector are constrained to be equal over the interfailure interval between the i-th and i + 1-th failure days (including the i + 1-th failure day): $\gamma_{t(i)+1} = \gamma_{t(i)+2} = \ldots = \gamma_{t(i+1)}$. This step model is a straightforward generalization of the simple constant hazard model we presented earlier. But the goal of the description here is primarily to show how nonparametric models fit into the bigger picture which we will be developing; we would generally not recommend that researchers use our SAS PROC NLMIXED approach to fit these nonparametric models. Very good special purpose software already exists that is perfectly satisfactory for fitting these models, or models that are close enough.

The impulse model corresponds to the KME or the generalized KME, or GKME. In modern usage the KME usually refers specifically to the version of Kaplan and Meier's (1958) estimator appropriate for untruncated data. As implemented in many programs such as SAS PROC LIFETEST, the KME does not allow for delayed entry (left-truncation). Hyde (1977) points out that a close reading of Kaplan and Meier (1958: 463, Eq. 2b) shows that they also explicitly treated left-truncation as well. Lynden-Bell (1971) appears to be the first to give a detailed consideration of nonparametric estimation of $S(t)$ in the presence of truncation (Woodroofe 1985), and presents the generalization of the KME, the GKME. The GKME has been reinvented numerous times from various perspectives; Pollock et al. (1989) popularized this estimator in wildlife telemetry studies.

As noted, Kaplan and Meier (1958) demonstrated that what they called the product limit estimator was the nonparametric maximum-likelihood estimator (NPMLE) of $S(t)$ for Case I observations. Although NPMLEs are of great theoretical interest, this does not imply that NPMLEs are in any sense best estimators. Nonparametric maximum likelihood is not the same thing as ordinary maximum likelihood. The optimality properties of ordinary maximum likelihood do not necessarily carry through to NPMLEs (Cox 1972, Anderson et al. 1992).

The step-hazard model is closely, and confusingly, related to another popular nonparametric survival estimator, the Breslow survival estimator. Indeed, the step-hazard model is sometimes called the Breslow hazard model. However, as Miller (1981) notes, Breslow (1974) extended his step-hazard structure to his survival estimator in a manner that does not appear to be consistent with equation (1), and the resulting Breslow survival estimator essentially appears to be based on an impulse-hazard model. Link (1984) fixed this, and developed a survival estimator that is directly consistent with Breslow's step-hazard model through equation (1); we will refer to this as the Breslow-Link model. We mention Breslow-Link only because it is the approach that is exactly consistent with our general development.

In practice GKME, Breslow, or Breslow-Link will usually give very similar answers, and no clear theoretical reason exists for preferring one over another if one has Case I or once-daily monitored Case II data. SAS PROC PHREG is a good software choice for either the GKME or the Breslow approach. We are not aware of an implementation of Breslow-Link, but either GKME or Breslow are fine substitutes. To accommodate the left-truncation, that is,

entry after age $t = 0$, one must use the ENTRY = varname model statement option, where varname is the SAS variable giving the age at which the nest was discovered. Using a KME procedure such as SAS PROC LIFETEST that assumes entry at age $t = 0$ will result in a potentially biased results because early failures will be underrepresented (Tsai et al. 1987), much like the apparent estimator of nest success is biased. To obtain survival estimates in PROC PHREG, one specifies a null model without any covariates and includes a BASELINE statement. One can specify either the GKME model with the BASELINE METHOD = PL or the Breslow approach with BASELINE METHOD = CH.

Because of the requirement of continuous or near continuous monitoring, these procedures cannot be recommended for application to our general Case II example data. GKME or Breslow are not appropriate because the exact day of failure is not known due to interval-censoring. In addition, KME is not appropriate because it ignores the left-truncation. However, we applied these techniques to examine the consequences. For these analyses, if a failure was observed, we used the midpoint of the failure interval as the exact age at which the failure occurred. We used SAS PROC PHREG to obtain KME (Program B-1, Appendix 2) and GKME (Program B-2, Appendix 2) estimates. By not including the ENTRY statement, the resulting KME assumes all nests are discovered at age 0, (nest initiation), and as expected, this resulted in a substantial upward bias in the estimated survival curve (solid circles, Fig. 3). The GKME (hollow diamonds, Fig. 3) correctly accommodates the left-truncation (delayed entry), but the midpoint assumption appears to cause bias at the youngest ages because the relative long initial intervals prevent any imputed failure times near initiation. By the end of the nesting period, the GKME is not too dissimilar from the more appropriate estimators presented later. The problems observed with the KME and GKME are predictable consequences of the incorrectly specified likelihood macrostructures.

General Case II—Nonparametric Survival Estimation

Turnbull (1976) developed the general theory for obtaining NPMLE's of $S(t)$ for interval-censored and truncated data. Pan and Chappell (1999) later showed that Turnbull's estimator would not always work when the data are sparse, and provided a correction. Even when this approach works in the sense of giving consistent estimates, the estimates may be unstable (Lindsey and Ryan 1998). Generally speaking, Turnbull's and related NPMLE algorithms are seeking the points at which the hazard should have impulses similar to GKME. The goal of nonparametric maximum likelihood estimation is to find the maximum number of impulses that can be estimated, but this means the problem often teeters on the brink of over-parameterization. In the real world, it is usually unlikely that the hazard function swings wildly up and down from day to day (except from known events such as storms that can be accounted for), and the flexibility of a fully nonparametric estimator is, in general, wasted. By imposing a minimal amount of structure on the daily hazard rates, we can avoid the problems with instability yet still maintain flexibility. We explore this idea of weakly structured models next.

General Case II—Weakly Structured Survival Estimation

The simple solution to the problems of a fully nonparametric approach is to use the step-hazard model with fewer than the maximum number of possible steps, which preserves flexibility yet permits reliable estimation. This is an easy extension of the simple constant-hazard model $h(t) = \lambda$ we presented previously. We now break the time line into intervals at our discretion, and if age t falls into the κ-th interval, we have:

$$h(t) = \lambda_{\kappa}$$

which constrains all of the Λ_t's (or γ_t's) in interval k to be equal.

This form of the step-hazard model has a long history in biostatistics as a convenient weakly structured survival model (Oakes 1972; Holford 1976, 1980; Laird and Oliver 1980, Anderson et al. 1997, Kim 1997, Lindsey and Ryan 1998, Ibrahim et al. 2001), and it is the logical companion of the Breslow-Link nonparametric model. It has been referred to as semi-parametric (Laird and Oliver 1980) or loosely parametric (Cai and Betensky 2003). This model adapts well to interval-censored data (Kim 1997, Lindsey and Ryan 1998), who both present EM (expectation-maximization) algorithms for estimation in the untruncated setting. However, in our experience Newton-type maximization algorithms such as used by SAS PROC NLMIXED work fine as long as starting values are selected carefully. An effective strategy for step or piecewise models is to fit models with progressively more pieces, using the previous estimates as starting values in an obvious way. Lindsey and Ryan (1998) discuss strategies for positioning the steps.

We applied this approach to our example data with steps somewhat arbitrarily placed

every 5 d (Program A-2, Appendix 1). The results suggest some irregularity in the age-specific survival, with a perhaps an inflection around day 15 (crosses in Fig. 3).

General Case II—Parametric Survival Estimation

We have already considered the simplest hazard model $h(t) = \lambda$, the constant or age-independent model which results in exponentially distributed failure times. In biostatistical survival analyses, many other popular parametric-hazard models correspond to different ideas about how the hazards change with age. An especially popular one is the Weibull (Kalbfleisch and Prentice 1980). The hazard function for the Weibull is given as $h(t) = \lambda\rho(\lambda t)^{\rho-1}$, which allows the failure hazard to change smoothly with age, either increasing or decreasing depending on the parameter ρ (the Weibull reduces to the exponential model when $\rho = 1$). Because our NLMIXED approach is based in the daily cumulative hazard rather than the hazard $h(t)$ directly, we need the daily cumulative hazard to obtain exact maximum likelihoods, which after a simple integration is found to be $\Lambda_t = \lambda^\rho[(t)^\rho - (t - 1)^\rho]$ (Kalbfleisch and Prentice (1980). In terms of γ_t, we have $\gamma_t = \rho\varphi + \log(t^\rho - (t-1)^\rho)$, where $\varphi = \log(\lambda)$ (Program A-3, Appendix 1). Figure 3 shows the Weibull fit to the example data (hollow squares) drops away more rapidly than the exponential model, and generally produces the lowest survival estimates of any of the procedures. In this example, the weakly structure estimates are bracketed by the exponential and Weibull although there is no reason to expect this in general. The Weibull shape parameter ρ was estimated to be 0.80 with 95% confidence intervals of 0.51–1.10, so on this basis it cannot be claimed that the Weibull is a significant improvement over the exponential. Indeed, as measured by Akaike's information criterion (AIC) (Burnham and Anderson 2002), the exponential model (AIC = 594.1) is as good as or better than the Weibull (AIC = 594.4) and better than the weakly structured model (AIC = 601.4). Some would no doubt argue that this shows the potential advantages of parametric models (Miller 1983), while others might not (Meier et al. 2004). At least in our example, it does not appear to matter much which hazard model is used but this of course cannot be counted on in general.

Many other parametric hazard models have been proposed (Kalbfleisch and Prentice 1980). Sometimes these are justified on the basis of some underlying theory that gives rise to their particular form, but they are frequently used in a less theoretical curve-fitting mode. For pure curve fitting, one could postulate a quadratic trend by specifying a hazard function $h(t) = \exp(a + bt + ct^2)$. With a little more programming, this curve-fitting approach could be extended to very flexible models such as polynomial splines (i.e., piecewise polynomial models that satisfy certain continuity constraints at the knots that join them). The most basic such piecewise polynomial spline model is the step-function model discussed previously.

If using parametric survival-analysis software such as SAS PROC LIFEREG, one must be careful that both the interval-censoring and left-truncation are appropriately handled. For example, LIFEREG can accommodate interval-censoring but not left-truncation. As with KME, ignoring left-truncation in parametric models can seriously bias survival estimates upward.

General Case II—Regression Analysis

Proportional Hazards Analysis of Covariates

Within the above framework, regression analyses are easy. Let $\boldsymbol{X}$ be a row vector of covariates, and let β be a column vector of regression coefficients. The log-hazard function $ln\ (h(t))$ can assume any value from $-\infty$ to ∞, so it is natural to model it with a typical linear model $ln\ (h(t \mid \boldsymbol{X})) = \beta_0(t) + \boldsymbol{X}\beta$. This can also be expressed as the multiplicative model $h(t \mid \boldsymbol{X}) = h_0(t)\exp(\boldsymbol{X}\beta)$ which is the proportional-hazards (PH) model popularized by Cox (1972). The covariate-specific term $\exp(X_i\beta_i)$ is the hazard ratio, and scales the hazard function up or down. The unit hazard ratio $\exp(\beta_i)$ indicates how much a unit shift in X_i shifts the hazard function.

The baseline hazard function $h_0(t)$ is the value $h(t \mid \boldsymbol{X})$ assumes when all covariate values are 0 (when $\boldsymbol{X} = \boldsymbol{0}$, $\exp(\boldsymbol{X}\beta) = 1$). Under the proportional-hazards assumption, we have the relationship $ln\ \Lambda_t(\boldsymbol{X}) = \gamma_{0t} + \boldsymbol{X}\beta$, where the intercept γ_{0t} is the log baseline cumulative daily hazard. Covariates are easily included in any of the analyses illustrated above simply by adding $\boldsymbol{X}\beta$ to each element of the vector gamma.

The models presented here are essentially generalizations of Prentice and Gloeckler's (1978) grouped data PH model, generalized for left-truncation and overlapping intervals. Very useful background can found in Section 4.6 of Kalbfleisch and Prentice (1980). Our approach extends Lindsey and Ryan's (1998) piecewise treatment of interval-censored data to left-truncated data as well. When the above regression approach is applied to Case I or once-daily monitored Case II data, the result is the full-likelihood version of the Cox

model. Cox invented the idea of partial likelihood, in which one can essentially ignore all of the likelihood except that portion that contains the covariates and their coefficients and thus avoid estimating the γ_t's. This has great computational benefits for large data sets but otherwise no reason is evident to prefer partial maximum-likelihood estimates. For Case I or once-daily monitored Case II data, it will generally be more convenient to use commercial software (e.g., SAS PROC PHREG) that accommodates delayed entry. However, we are not aware of a commercial program that correctly accommodates general left-truncated, interval-censored data that are typical of many nest-survival studies.

Alternative Regression Approaches (Advanced)

In addition to PH models, accelerated failure time (AFT) models and proportional discrete hazards odds (PDHO) models enjoy some popularity in survival analysis. AFT models that allow weakly structured modeling of the baseline have not been well developed and we will not consider them further. PDHO models can be traced to at least Cox's original 1972 paper; they are best suited to situations where the failure events are occurring in truly discrete time (Breslow 1974, Thompson 1977, Kalbfleisch and Prentice 1980: Eq. 2.23.). Truly discrete time-failure processes are relatively rare in nature, and require the event probability to be zero at almost all times except a countable number of instances. An example of a truly discrete time failure process is the repeated slamming of a car door in reliability testing (B. Storer, pers. comm.)

For example, assume that all failed nests fail at an instant before the end of the monitoring day. Then, the daily mortality probability for day t, $M(t \mid t-1)$ places all its probability mass at that single instant, which we will call $\delta_t = M(t \mid t-1)$, the discrete hazard function. In proportional daily discrete hazards odds (PDDHO) models, the daily odds

$$\theta_t(\boldsymbol{X}) = \frac{\delta_t(\boldsymbol{X})}{1-\delta_t(\boldsymbol{X})}$$

takes the place of the cumulative daily hazard $\Lambda_t(\boldsymbol{X})$ in PH models. The log PDDHO model is then $ln\ \theta_t(\boldsymbol{X}) = \alpha_{0t} + \boldsymbol{X}\alpha$, where

$$\alpha_{0t} = \ln\left(\frac{\delta_t(0)}{1-\delta_t(0)}\right)$$

and α is the vector of log odds ratios. This posits a logistic regression model for daily failures. In terms of log daily cumulative hazards, the PDDHO model can be expressed as $\gamma_t = \log(\log(1 + \exp(\alpha_{0t} + \boldsymbol{X}\alpha)))$, which allows us to fit PDDHO models within our general hazards framework. When daily survival is moderately high, the PH and PDDHO will return similar results in most survival applications as long as the likelihood macrostructure is correctly represented (Thompson 1977). Efron (1988) illustrates the application of the PDHO model in what is essentially a once-monthly monitoring situation and relates it back to hazard functions. The approaches of Dinsmore et al. (2002), Rotella, at al. (2004), and Shaffer (2004a) are examples of general Case II nest-survival analyses with correctly specified PDDHO models. Given the similarity of results in most cases, the primary reason for preferring the PH approach over PDHO are theoretical rather than practical. The PDHO model for grouped data assumes that one has discovered the time interval at which the survival process acts in a proportional odds manner. If a process follows a PDHO process for a daily interval, it cannot obey a PDHO process for any other interval width and hence the interpretation of the regression coefficients α depends in the interval choice. The PH approach is interval invariant; $h(t \mid \boldsymbol{X}) = h_0(t)\exp(\boldsymbol{X}\beta)$, $\Lambda_t(\boldsymbol{X}) = \Lambda_t(\boldsymbol{0})\exp(\boldsymbol{X}\beta)$, and $S(t \mid \boldsymbol{X}) = S(t \mid \boldsymbol{X} = \boldsymbol{0})^{\exp(\boldsymbol{X}\beta)}$ are all equivalent representations of the PH model.

General Case II—Regression Example

For our example data set, nests in the sample were obtained by searching right-of-way habitat along Interstate 94 in south-central North Dakota. We examined whether distance to the road shoulder was associated with survival (Programs A-4, A-5, A-6; Appendix 1); the unit of distance measurement was meters. These data are summarized in Table 2 of Shaffer (2004a). Generally speaking, the effect of model misspecification in the regression analysis of survival data is to weaken the covariate association and that indeed appears to be consistent with what we observe (Table 1). The three models with correctly specified macro-structures give similar results regardless what hazard structure (constant, Weibull, step) was assumed, although increasing the flexibility of the baseline appears to slightly increase the variance (decrease the t-ratio). A hazard ratio of 1.016 means that for every meter away from the shoulder, the failure hazard $h(t)$ or $\Lambda(t)$ increases by a factor of 1.016. Thus, X meters from the shoulder the hazard ratio is $H(X) = 1.016^X$. In terms of age-specific survival, this means the survival of a nest distance X meters from the shoulder is $S(t \mid X) = S(t \mid X = 0)^{H(X)}$, where $S(t \mid X = 0)$ is the survival

TABLE 1. HAZARD AND ODDS RATIOS FOR MODELS FITTED TO THE BLUE-WINGED TEAL (*ANAS DISCORS*) DATA.

Model	Hazard ratio (*t*) [a]	Odds ratio (*t*)
Constant hazard (or odds)	1.016 (2.00)	1.016 (2.00)
Weibull hazard	1.016 (1.99)	–
Step-hazard (or odds)	1.015 (1.91)	1.016 (1.91)
Cox/GKME baseline	1.014 (1.76)	–
Cox/KME baseline	1.012 (1.52)	–

[a] The number in parentheses is the t-ratio for the log-hazard ratio: estimate/(SE).

immediately at the shoulder. The Cox-GKME approach (Program B-3, Appendix 2) fails to model the interval censoring, and results in a somewhat weakened covariate association. The Cox-KME (Program B-4, Appendix 2) approach which fails to model both the left-truncation and interval-censoring results in an even weaker association. No appreciable difference occurs between the hazard-ratio (PH) or odds-ratio (PDDHO) formulation (Programs A-7, A-8; Appendix 2). PDDHO models can be cast equally well in terms of mortality odds as we have done or survival odds as Shaffer (2004a) did, which accounts for why his log odds ratio for this example is the same as ours except for the sign.

TIME AND TIME-VARYING COVARIATES AND COEFFICIENTS (ADVANCED)

So far, the most general regression model we have considered is:

$$h(t \mid \boldsymbol{X}) = h_0(t)\exp(\boldsymbol{X}\beta),$$

where t is age. However, in its fullest generality we can have

$$h(t,c \mid \boldsymbol{X}(t,c)) = h_0(t,c)\exp(\boldsymbol{X}(t,c)\,\beta(t,c)),$$

where c refers to calendar time. This model incorporates three new features: (1) a bivariate calendar time/age baseline hazard function, (2) time and/or age varying covariates, and (3) time and/or age varying coefficients. We will describe each of these briefly. For sticklers, we note that we are appealing here to the mean value theorem for integrals to justify blurring the distinction between $h(t)$ and Λ_t, and we avoid the complication of integrating $h(t,c \mid \boldsymbol{X}(t,c))$ out over the day $t-1$ to t.

Bivariate time/age baseline

Before, we constructed a piecewise step function for the age-specific hazard. We can take a similar approach for calendar time. This can be thought of as dividing the Lexus diagram into a patchwork of rectangles. Let k index the age intervals, and let m index the time intervals. Then for the resulting rectangle indexed by km, we can posit the log daily cumulative-hazard model $\gamma_k + \tau_m$. This log-linear model implies conditional independence of age and time (Bishop et al. 1975), as the daily cumulative hazard for each day is the product of a day term and a time term. An age-time interaction model is constructed by defining an individual term for each rectangle km. For this weakly structured age-time approach to work well, one must be judicious with respect to the number and position of the rectangles.

Time and age varying covariates

It is fairly easy to build time or age-varying covariates into the generic SAS PROC NLMIXED approach by using arrays that allow the covariate values to change as age or time changes. The use and interpretation of time-varying covariates requires care. Kalbfleisch and Prentice (1980) identify two general classes of time-varying covariates—external and internal. An internal covariate is something measured from the nest, such as the number of eggs or presence of parasitism and depends on the existence of the nest to be measured. As the name implies, an external covariate is one measured external to the nest, such as temperature or rainfall. Internal time-varying covariates are problematic with interval monitoring because the covariate values themselves will be interval-censored. The most common approach is to take the most recent value forward in time, although this is not without issues (Do 2002). Interpreting internal time-varying covariates can be problematic. For example, if parasitism is associated with nest failure, it is difficult to conclude directly whether parasitism is causal or simply associated with frail nests predisposed to fail regardless.

Even for a fixed covariate such as distance to the road, say X, we may be interested in whether its effect changes with age or time. We can model this as $(\alpha + \beta t)X$, where $\alpha + \beta t$ is viewed as a generalized regression coefficient of X that is a linear function of age t. We applied this to our example data using the weakly structured baseline model (Program A-9, Appendix 1);

no suggestion arose that the road effect varied with age (t-ratio = –0.24). Of course more flexible age-varying models could be specified as well. At the highest level of generality, one can have time/age-varying covariates with time/age-varying coefficients.

Frailty (Random Effects) and Spatial Models (Advanced)

In addition to allowing traditional fixed-effect regression models, some programs such as SAS PROC NLMIXED allow the inclusion of random effects. Such models are appealing because they allow a mechanism for modeling nests reasonably expected to have correlated fates. For example, for nests near an ephemeral pond, the fates of all nests may share some statistical association, if the pond dries up. We could reflect this by adding a random pond effect in the proportional hazards model, where z_j is the random effect of pond j, giving the mixed model $ln\ \lambda_t(\boldsymbol{X}, j) = z_j + \gamma_t + \boldsymbol{X}\beta$.

Random effects in survival models require some special considerations. In survival-analysis, random-effects models such as just described are called shared-frailty models, with z_j being an unobserved frailty factor shared by all members in cluster j. Frailties have the effect of making the population (marginal) hazard decline over time because subjects with large frailties (large z_j) get eliminated first, and the remaining population becomes progressively shifted toward small z_j as time goes by. This is problematic in nest-survival studies because of left-truncation: the frailty distribution for discovered nests will be a function of the age of discovery as well as other covariates.

To clarify this, suppose it is possible to find all nests at the time of initiation. In this case, no nests would be overlooked, and we would be aware of all clusters. The typical assumption is that the cluster random effect z_j is normally distributed with mean 0 and variance σ^2, i.e., $N(N, \sigma^2)$. If the discovery of nests is delayed, some nests will fail and be unavailable for discovery. In some cases, all the nests in a cluster will fail so the cluster cannot even be identified. Because the initial z_j influences the likelihood that all nests in the cluster will be destroyed and later unavailable for discovery, the z_j of the discovered clusters are a biased sample from $N(0, \sigma^2)$, the mean of which will be shifted to the left toward the less frail. This will be most problematic in situations where some clusters have few nests initiated to begin with, and an especially troublesome scenario is when the random effect is associated with both the number of nests initiated in a cluster as well as survival in the cluster (i.e., birds should avoid nesting in habitat where success is likely to be low). Additional work is needed to better understand the practical significance of this issue and to develop strategies for addressing it.

Frailty models for left-truncated data have received relatively little attention in survival analysis (Huber-Carol and Vonta 2004, Jiang et al. 2005), and more work is needed before reasonable guidelines can be given on this. Natarajan and McCulloch (1999) present some models of heterogeneity for nest-survival data, but their approach appears to be difficult to relate to a standard hazards-based frailty approach. With the increasing interest in including spatial information into ecological analyses, this problem is especially urgent because spatial correlation in survival models is most conveniently accounted for with frailty models (Banerjee et al. 2003). Extending such analyses to left-truncated data is an important and challenging problem that should be a research priority.

Before leaving the topic of frailties, it is interesting to note their relationship with covariates. Suppose the failure process obeys the regression relationship:

$$ln\ (\Lambda_t) = \gamma + X,$$

where we assume the baseline γ does not depend on age and X is some continuous covariate. If we do not observe X and fit just a baseline model, we will observe that the baseline γ_t declines with age due to the frailty effect induced by X, despite the fact that an individual nest's hazard is not age-dependent. This points out the importance of allowing for flexible baselines as one explores different models.

ESTIMATION AND PREDICTION

We used the relationship

$$S(t) = \exp\left[- \sum_{i=1}^{t} \exp(\gamma_i) \right]$$

to obtain the estimates displayed on Fig. 3. The ESTIMATE statement in SAS PROC NLMIXED could be used to obtain standard errors as well. We now briefly consider what this is an estimate of, and what assumptions are involved. For the estimate of $S(t)$ to have meaning, the samples on which it was based must have been representative of some population of interest. The ideal situation would be to have a representative sample of all initiated nests, but delayed discovery and resulting left-truncation ensures this is usually unobtainable. But what we can hope for is that when we discover a nest at age r, it is representative of all initiated nests that

then survive to age r. If this condition is met, a correctly specified likelihood takes care of the left-truncation issues.

What might cause a nest discovered at age r not to be representative of all initiated nests that survive to age r? This can occur whenever the discovery of active nests is also associated with covariates that affect survival. For example, suppose active nests are more easily discovered close to water, and suppose independently of this, nests close to the water have higher survival. Such enhanced discovery will bias the number of close water nests in the sample above and beyond the bias caused by their higher survivability alone. The result will be that the estimate of $S(t)$ is in turn biased high and not representative of all initiated nests.

On the other hand, the regression evaluation of covariates does not require that the sample be representative of the active nests and indeed sample collection may attempt to disproportionately obtain nests with particular covariate values for increased power.

This emphasizes the importance of carefully planned sampling designs that weigh the various goals of survival estimation versus covariate assessment.

A goal closely related to that of estimation is that of prediction. That is, if we observed that cover density, say X, is associated with nest survival, it would be interesting to predict how overall survival would respond if X were manipulated. This is a nontrivial problem, and involves estimating the distribution of X associated with the nests at the time of initiation. This problem is considered by Shaffer and Thompson (*this volume*). Extending these considerations to random effects models, which involves integrating over the random effects distribution, seems especially challenging.

DISCUSSION

Our primary goal was to embed nest survival into the biostatistical approach to survival analysis. This provides both a sound theoretical foundation as well as a large toolbox from which to choose techniques. Such a unified framework permits judging the strengths and weaknesses of recently proposed nest survival techniques, such as the logistic-exposure model (Shaffer 2004a) or Kaplan-Meier and Cox applications (Nur et al. 2004). From basic survival-analysis considerations, we propose a new class of nest-survival analyses based on the complementary log-log link function. This framework is well-suited for use with weakly structured hazard models, which combine the flexibility of nonparametric models with the stability of fully parametric procedures.

Given their immense popularity in human biostatistics, some readers may be surprised that we did not devote more attention to fully nonparametric procedures. Fully nonparametric approaches work remarkably well for untruncated and right-censored data (Meier et al. 2004), but the resulting enthusiasm should not be automatically conferred to the left-truncated and interval-censored situation. Indeed, unless at least a few nests are discovered on the day of initiation, left-truncation will even prevent the fully nonparametric estimation of the survival function. Weakly structured approaches, while not a panacea, ameliorate these problems to a large extent.

Many weakly structured procedures, including those presented here, can be thought of as attempts to approximate the hazard function with a piecewise polynomial spline function. Piecewise models such as we presented are the simplest example, and constitute a 0-order B-spline basis. Smoother approximations can be obtained by specifying more complex splines, but this comes at the cost of additional parameters to estimate. A very appealing solution would be to employ a penalized spline approach (Gray 1992, Cai and Betensky 2003), but software is unavailable.

Although some theoretical holes still exist (e.g., frailty models), in general nest-survival theory has progressed well beyond the readily available software. It would be nice to be able to avoid the arbitrariness of the piecewise hazard approach with either an optimally smoothed spline (Gray 1992, Heisey and Foong 1998) or Bayesian approach (He et al. 2001, He 2003), but user-friendly software that includes regression analysis is not yet available. Theoretical and practical work is needed to extend the ideas of model goodness-of-fit and residuals from the continuous monitoring situation (Therneau and Grambsch 2000) to interval-censoring. User-friendly software which would allow covariate analysis of both survival and discovery probabilities is needed for the general Case III situation (Heisey 1991).

ACKNOWLEDGMENTS

Special thanks are due to Stephanie Jones, who helped improve both the substance and form of this paper. Christine Bunck, Bobby Cox, Ken Gerow, and an anonymous reviewer provided many helpful comments and suggestions. Douglas Johnson and the late Albert T. Klett collected the data used in our examples.

APPENDIX 1. INTERVAL-CENSORED EXAMPLES.

```
libname local '';
options ls=75 ps=50;
data a;
  set local.bwteal;
  run;

/*
Variables in the data set are:
nestid   (nest id)
firstday (age on first day of interval)
lastday  (age on last day of interval)
success  (whether interval was survived(1) or not(0))
d2road   (covariate; distance to road)
*/

/* Basic macro used by all methods; corresponds to pseudo-code in text */

%MACRO CASE2ML;
  PROC NLMIXED DATA=A DF=99999;
   %INITPARM;
   ARRAY GAMMA {*} X1-X35;
   %GAMMAMOD;
   CUMHAZ = 0;
   DO DAY = firstday to lastday-1;
     DAYCUMHZ = EXP(GAMMA[DAY]);
     CUMHAZ = CUMHAZ + DAYCUMHZ;
   END;
   SURVIVE = EXP(-CUMHAZ);
   MODEL success~BINARY(SURVIVE);
   %ESTIMATE;
   RUN;
%MEND;

/* ------------------------------------------------------------ */

TITLE 'PROGRAM A-1: Constant hazard; Johnson-Bart-Robson model';

%MACRO INITPARM;
  PARMS g1=-3.3;
%MEND;

%MACRO GAMMAMOD;
  DO AGE = 1 TO 35;
   GAMMA [AGE] = g1;
  END;
%MEND;

%MACRO ESTIMATE;
  ESTIMATE 'DSR' EXP(-EXP(g1));
%MEND;

%CASE2ML;

/* ------------------------------------------------------------ */

TITLE 'PROGRAM A-2: Piecewise constant hazard; weakly structured';

%MACRO INITPARM;
  PARMS g1=-3 g2=-3 g3=-3 g4=-3 g5=-3 g6=-3 g7=-3;
%MEND;

%MACRO GAMMAMOD;
  DO AGE = 1 TO 35;
```

```
  IF   (AGE LE 5)    THEN GAMMA [AGE] = g1;
  ELSE IF(AGE LE 10) THEN GAMMA [AGE] = g2;
  ELSE IF(AGE LE 15) THEN GAMMA [AGE] = g3;
  ELSE IF(AGE LE 20) THEN GAMMA [AGE] = g4;
  ELSE IF(AGE LE 25) THEN GAMMA [AGE] = g5;
  ELSE IF(AGE LE 30) THEN GAMMA [AGE] = g6;
  ELSE GAMMA [AGE] = g7;
  END;
%MEND;

%MACRO ESTIMATE;
  ESTIMATE 'DAILY BASELINE, INTERVAL 1' EXP (-EXP (g1));
  ESTIMATE 'DAILY BASELINE, INTERVAL 2' EXP (-EXP (g2));
  ESTIMATE 'DAILY BASELINE, INTERVAL 3' EXP (-EXP (g3));
  ESTIMATE 'DAILY BASELINE, INTERVAL 4' EXP (-EXP (g4));
  ESTIMATE 'DAILY BASELINE, INTERVAL 5' EXP (-EXP (g5));
  ESTIMATE 'DAILY BASELINE, INTERVAL 6' EXP (-EXP (g6));
  ESTIMATE 'DAILY BASELINE, INTERVAL 7' EXP (-EXP (g7));
%MEND;

%CASE2ML;

/* ------------------------------------------------------------ */

TITLE 'PROGRAM A-3: Weibull hazard';

%MACRO INITPARM;
  PARMS rho=1 loglam=-3;
%MEND;

%MACRO GAMMAMOD;
  GAMMA [1] = rho*loglam + LOG(1);
  DO AGE = 2 TO 35;
  GAMMA [AGE] = rho*loglam + LOG(AGE**rho - (AGE-1)**rho);
  END;
%MEND;

%MACRO ESTIMATE;
%MEND;

%CASE2ML;

/* ------------------------------------------------------------ */

TITLE 'PROGRAM A-4: Constant hazard with covariate';

%MACRO INITPARM;
  PARMS g1=-3.3 beta=0;
%MEND;

%MACRO GAMMAMOD;
  DO AGE = 1 TO 35;
   GAMMA [AGE] = g1 + beta*d2road;
  END;
%MEND;

%MACRO ESTIMATE;
  ESTIMATE 'Hazard Ratio' EXP(beta);
%MEND;

%CASE2ML;

/* ------------------------------------------------------------ */
```

```
TITLE 'PROGRAM A-5: Piecewise constant hazard with covariate';

%MACRO INITPARM;
  PARMS g1=-3 g2=-3 g3=-3 g4=-3 g5=-3 g6=-3 g7=-3 beta=0;
%MEND;

%MACRO GAMMAMOD;
  DO AGE = 1 TO 35;
  IF   (AGE LE 5)    THEN GAMMA [AGE] = g1;
  ELSE IF(AGE LE 10) THEN GAMMA [AGE] = g2;
  ELSE IF(AGE LE 15) THEN GAMMA [AGE] = g3;
  ELSE IF(AGE LE 20) THEN GAMMA [AGE] = g4;
  ELSE IF(AGE LE 25) THEN GAMMA [AGE] = g5;
  ELSE IF(AGE LE 30) THEN GAMMA [AGE] = g6;
  ELSE GAMMA [AGE] = g7;
  GAMMA [AGE] = GAMMA [AGE] + beta*d2road;
  END;
%MEND;

%MACRO ESTIMATE;
  ESTIMATE 'Hazard Ratio' EXP(beta);
%MEND;

%CASE2ML;

/* ------------------------------------------------------------ */

TITLE 'PROGRAM A-6: Weibull hazard with covariate';

%MACRO INITPARM;
  PARMS rho=1 loglam=-3 beta=0;
%MEND;

%MACRO GAMMAMOD;
  GAMMA [1] = rho*loglam + LOG(1) + beta*d2road;
  DO AGE = 2 TO 35;
  GAMMA [AGE] = rho*loglam + LOG(AGE**rho - (AGE-1)**rho) + beta*d2road;
  END;
%MEND;

%MACRO ESTIMATE;
  ESTIMATE 'Hazard Ratio' EXP(beta);
%MEND;

%CASE2ML;

/* ------------------------------------------------------------ */

TITLE 'PROGRAM A-7: Constant odds with covariate';

%MACRO INITPARM;
  PARMS t1=-3.3 alpha=0;
%MEND;

%MACRO GAMMAMOD;
  DO AGE = 1 TO 35;
   GAMMA [AGE] = log(log(1 + exp(t1 + alpha*d2road)));
  END;
%MEND;

%MACRO ESTIMATE;
  ESTIMATE 'Odds Ratio' EXP(alpha);
%MEND;

%CASE2ML;

/* ------------------------------------------------------------ */
```

```
TITLE 'PROGRAM A-8: Piecewise constant odds with covariate';

%MACRO INITPARM;
  PARMS t1=-3 t2=-3 t3=-3 t4=-3 t5=-3 t6=-3 t7=-3 alpha=0;
%MEND;

%MACRO GAMMAMOD;
  DO AGE = 1 TO 35;
  IF   (AGE LE 5)    THEN GAMMA [AGE] = log(log(1 + exp(t1 + alpha*d2road)));
  ELSE IF(AGE LE 10) THEN GAMMA [AGE] = log(log(1 + exp(t2 + alpha*d2road)));
  ELSE IF(AGE LE 15) THEN GAMMA [AGE] = log(log(1 + exp(t3 + alpha*d2road)));
  ELSE IF(AGE LE 20) THEN GAMMA [AGE] = log(log(1 + exp(t4 + alpha*d2road)));
  ELSE IF(AGE LE 25) THEN GAMMA [AGE] = log(log(1 + exp(t5 + alpha*d2road)));
  ELSE IF(AGE LE 30) THEN GAMMA [AGE] = log(log(1 + exp(t6 + alpha*d2road)));
  ELSE GAMMA [AGE] = log(log(1 + exp(t7 + alpha*d2road)));
  END;
%MEND;

%MACRO ESTIMATE;
  ESTIMATE 'Odds Ratio' EXP(alpha);
%MEND;

%CASE2ML;

/* ------------------------------------------------------------ */

TITLE 'PROGRAM A-9: Piecewise constant hazard with covariate';

%MACRO INITPARM;
  PARMS g1=-3 g2=-3 g3=-3 g4=-3 g5=-3 g6=-3 g7=-3 alpha=0 beta=0;
%MEND;

%MACRO GAMMAMOD;
  DO AGE = 1 TO 35;
  IF   (AGE LE 5)    THEN GAMMA [AGE] = g1;
  ELSE IF(AGE LE 10) THEN GAMMA [AGE] = g2;
  ELSE IF(AGE LE 15) THEN GAMMA [AGE] = g3;
  ELSE IF(AGE LE 20) THEN GAMMA [AGE] = g4;
  ELSE IF(AGE LE 25) THEN GAMMA [AGE] = g5;
  ELSE IF(AGE LE 30) THEN GAMMA [AGE] = g6;
  ELSE GAMMA [AGE] = g7;
  GAMMA [AGE] = GAMMA [AGE] + (alpha + beta * (AGE-15))*d2road;
  END;
%MEND;

%MACRO ESTIMATE;
%MEND;

%CASE2ML;
```

APPENDIX 2. KAPLAN-MEIER AND COX MODEL EXAMPLES.

```
libname local '';
options ls=75 ps=50;
data a;
  set local.bwteal;
  run;

/*
Variables in the data set are:
nestid   (nest id)
firstday (age on first day of interval)
lastday  (age on last day of interval)
success  (whether interval was survived(1) or not(0))
d2road   (covariate; distance to road)
*/

PROC SORT; BY nestid firstday;

DATA onerec;
  SET a;
  RETAIN entry;
  BY nestid firstday;
  IF first.nestid THEN entry = firstday - 1; /* visits at start of day */
  IF last.nestid THEN OUTPUT;

DATA onerec;
  SET onerec;
  IF success THEN time = lastday - 1;
  ELSE time = (firstday + lastday)/2 - 1;
  RUN;

TITLE 'Program B-1: KME model';
PROC PHREG data=onerec;
  MODEL time * success(1)=/;
  BASELINE OUT=out2 SURVIVAL=s2;
  RUN;

TITLE 'Program B-2: GKME model';
PROC PHREG data=onerec;
  MODEL time * success(1)=/ENTRY=entry;
  BASELINE OUT=out1 SURVIVAL=s1;
  RUN;

TITLE 'Program B-3: GKME model with covariate';
PROC PHREG data=onerec;
  MODEL time * success(1)=d2road/ENTRY=entry;
  RUN;

TITLE 'Program B-4: KME model with covariate';
PROC PHREG data=onerec;
  MODEL time * success(1)=d2road;
  RUN;
```